JOURNAL OF COMPLEXITY 2, 317–322 (1986)

On the Existence of Optimal Affine Methods for Approximating Linear Functionals

ALEKSEI G. SUKHAREV

*Faculty of Computational Mathematics and Cybernetics, Moscow State University, Moscow 119899, USSR**

Received April 15, 1986

The existence of an optimal affine method using linear information is established for the approximation of a linear functional on a convex set. This is a generalization of a result of S. A. Smolyak ("On Optimal Restoration of Functions and Functionals of Them," Candidate Dissertation, Moscow State University, 1965).

Let F be a set in a linear space and $S, x_1, \ldots, x_N$ functionals on F. Knowing that $f \in F$ and knowing the values $x_1(f), \ldots, x_N(f)$ we wish to construct an optimal method for approximating $S(f)$, i.e., a function θ_0 minimizing

$$\sup_{f\in F} |S(f) - \theta(x_1(f), \ldots, x_N(f))|$$

over the set of all real-valued functions θ. It is quite natural to look for a simple θ_0, and this is the case if θ_0 is a linear or affine function. It turns out that for convex F and linear $S, x_1, \ldots, x_N$ an optimal affine θ_0 exists.

THEOREM. *Let F be a convex set in a linear space and let $S, x_1, \ldots, x_N$ be linear functionals. Then there exists an affine function $\theta_0(y_1, \ldots, y_N) = p_0 + p_1 y_1 + \cdots + p_N y_N$ such that*

$$\sup_{f\in F}|S(f) - \theta_0(x_1(f), \ldots, x_N(f))| = \min_{\theta\in\Theta} \sup_{f\in F} |S(f) - \theta(x_1(f), \ldots, x_N(f))|,$$

where Θ is the set of all real-valued functions θ of variables

* Part of this work was done while the author was visiting the Department of Computer Science, State University of New York at Albany.

0885-064X/86 $3.00

$(y_1, \ldots, y_N) \in \{(y_1, \ldots, y_N) \mid y_1 = x_1(f), \ldots, y_N = x_N(f), f \in F\}.$

If the set F is balanced (i.e., $f \in F \Rightarrow -f \in F$), $p_0 = 0$.

This theorem was proven by Smolyak (1965) for continuous S, $x_1, \ldots,$ x_N (see also Bakhvalov (1971), where the proof is given for a balanced F). Further developments on this subject are due to Marchuk and Osipenko (1975), Osipenko (1976), Micchelli and Rivlin (1977), Traub and Woźniakowski (1980), Packel (1984), Werschulz and Woźniakowski (1985), and others. See Woźniakowski (1985) for a history of the subject.

To prove the theorem, we need the following well-known and almost obvious result.

LEMMA. *Let X, Y be arbitrary sets*

$$g: X \times Y \to R, \qquad \Phi = \{\varphi \mid \varphi: Y \to X\},$$

where Φ contains all mappings from Y to X. Define g on $\Phi \times Y$ by the formula $g(\varphi, y) = g(\varphi(y), y)$. Then

$$\inf_{\varphi \in \Phi} \sup_{y \in Y} g(\varphi, y) = \sup_{y \in Y} \inf_{x \in X} g(x, y).$$

Proof of Theorem. Let

$$Y = \{y = (y_0, y_1, \ldots, y_N) \mid y_0 = S(f), y_1 = x_1(f), \ldots, y_N = x_N(f), f \in F\}.$$

Since F is convex and S, $x_1, \ldots, x_N$ are linear, the set $Y \subset R^{N+1}$ is convex. For an arbitrary $\theta \in \Theta$ we have

$$\begin{aligned}\sup_{f \in F} |S(f) - \theta(x_1(f), \ldots, x_N(f))| &= \sup_{(y_0, y_1, \ldots, y_N) \in Y} |y_0 - \theta(y_1, \ldots, y_N)| \\ &= \sup_{(y_1, \ldots, y_N) \in \pi(Y)} \ \sup_{y_0 \in \sigma(y_1, \ldots, y_N)} |y_0 - \theta(y_1, \ldots, y_N)|,\end{aligned}$$

where $\sigma(y_1, \ldots, y_N) = \{y_0 \mid (y_0, y_1, \ldots, y_N) \in Y\}$ is the interval with the extreme points

$$a(y_1, \ldots, y_N) = \inf_{(y_0, y_1, \ldots, y_N) \in Y} y_0, \qquad b(y_1, \ldots, y_N) = \sup_{(y_0, y_1, \ldots, y_N) \in Y} y_0,$$

$\pi(Y) = \{(y_1, \ldots, y_N) \mid \sigma(y_1, \ldots, y_N) \neq \emptyset\}$ is the projection of Y on the subspace of the variables $y_1, \ldots, y_N$. From this and the lemma we get

$$\begin{aligned}&\inf_{\theta \in \Theta} \sup_{f \in F} |S(f) - \theta(x_1(f), \ldots, x_N(f))| \\ &\quad = \inf_{\theta \in \Theta} \ \sup_{(y_1, \ldots, y_N) \in \pi(Y)} \ \sup_{y_0 \in \sigma(y_1, \ldots, y_N)} |y_0 - \theta(y_1, \ldots, y_N)|\end{aligned}$$

$$= \sup_{(y_1,\ldots,y_N)\in\pi(Y)} \inf_{r\in R} \sup_{y_0\in\sigma(y_1,\ldots,y_N)} |y_0 - r| \tag{1}$$
$$= \sup_{(y_1,\ldots,y_N)\in\pi(Y)} \inf_{r\in R} \max\{b(y_1, \ldots, y_N) - r, r - a(y_1, \ldots, y_N)\}$$
$$= \sup_{(y_1,\ldots,y_N)\in\pi(Y)} \tfrac{1}{2}(b(y_1, \ldots, y_N) - a(y_1, \ldots, y_N)) \overset{\text{def}}{=} d.$$

If $d = +\infty$, then any $\theta \in \Theta$ is optimal. For $d = 0$ the proof is simple. Let $0 < d < +\infty$. To substantiate the existence of an optimal affine method, it is sufficient to point out a function $\theta_0(y_1, \ldots, y_N) = p_0 + p_1 y_1 + \cdots + p_N y_N$ such that

$$\sup_{f\in F} |S(f) - \theta_0(x_1(f), \ldots, x_N(f))| \le d$$

or, in other words,

$$|y_0 - (p_0 + p_1 y_1 + \cdots + p_N y_N)| \le d \qquad \text{for any } (y_0, y_1, \ldots, y_N) \in Y. \tag{2}$$

Consider the sets

$$Y_1 = Y - (d, 0, \ldots, 0)$$

and

$$Y_2 = Y + (d, 0, \ldots, 0)$$

(Fig. 1). Let

$$y = (y_0, y_1, \ldots, y_N) \in \operatorname{ri} Y_1 \cap \operatorname{ri} Y_2$$

(as usual, ri A denotes the relative interior of the set A). Then

$$(y_0 + d, y_1, \ldots, y_N) \in \operatorname{ri} Y, \qquad (y_0 - d, y_1, \ldots, y_N) \in \operatorname{ri} Y.$$

The line passing through these points clearly belongs to the affine hull of the set Y. The points

$$(y_0 + d + \epsilon/2, y_1, \ldots, y_N), \qquad (y_0 - d - \epsilon/2, y_1, \ldots, y_N)$$

lie on the same line. Hence, for a small $\epsilon > 0$, they are elements of the set Y. Therefore,

$$b(y_1, \ldots, y_N) - a(y_1, \ldots, y_N) \ge (y_0 + d + \epsilon/2) - (y_0 - d - \epsilon/2) = 2d + \epsilon,$$

which contradicts the last equality in (1). Thus, ri $Y_1 \cap$ ri $Y_2 = \emptyset$.

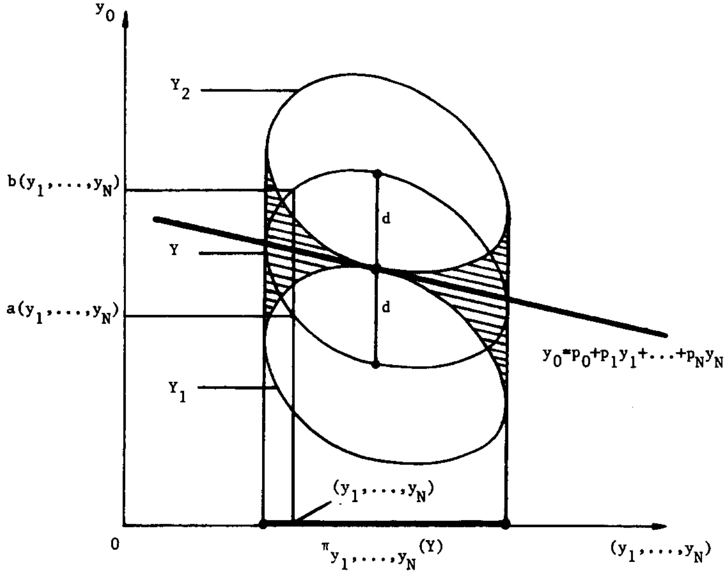

FIG. 1

Due to the separation theorem (see Rockafellar, 1970), in this case there exists a hyperplane $\langle c, y\rangle = \zeta$ separating the sets Y_1 and Y_2 properly, where $c = (c_0, c_1, \dots, c_N)$, $\langle c, y\rangle = \sum_{i=0}^{N} c_i y_i$. That is,

$$\langle c, y'\rangle \leq \zeta \leq \langle c, y''\rangle \qquad \text{for any } y' \in Y_1,\ y'' \in Y_2, \tag{3}$$

$$\langle c, \bar{y}'\rangle < \langle c, \bar{y}''\rangle \qquad \text{for some } \bar{y}' \in Y_1,\ \bar{y}'' \in Y_2. \tag{4}$$

Let $c_0 = 0$. Take any $y' = (y_0', y_1', \dots, y_N') \in Y_1$ and $y'' = (y_0' + 2d, y_1', \dots, y_N') \in Y_2$. We have $\langle c, y'\rangle = \langle c, y''\rangle$. Therefore, due to (3), $\langle c, y'\rangle = \zeta$ for any $y' \in Y_1$. Similarly, $\langle c, y''\rangle = \zeta$ for any $y'' \in Y_2$, which contradicts (4). The contradiction shows that $c_0 \neq 0$. Without loss of generality, let $c_0 > 0$. Define

$$p_1 = -c_1/c_0, \dots, p_N = -c_N/c_0, \qquad p_0 = \zeta/c_0.$$

Using the definition of Y_1 and Y_2, rewrite (3) in the form

$$y_0 - d - \sum_{i=1}^{N} p_i y_i \leq p_0 \leq y_0 + d - \sum_{i=1}^{N} p_i y_i$$

for any $y = (y_0, y_1, \dots, y_N) \in Y$, which verifies (2).

To complete the proof, we should show that $p_0 = 0$ for a balanced F. In this case, Y is also balanced.

Let $(y_1, \ldots, y_N) \in \pi(Y)$. Then, the points $(a(y_1, \ldots, y_N), y_1, \ldots, y_N)$ and $(b(y_1, \ldots, y_N), y_1, \ldots, y_N)$ belong to the closure $\bar{Y}$ of the set Y. Clearly, the set $\bar{Y}$ is balanced and convex. Hence, $(-a(y_1, \ldots, y_N), -y_1, \ldots, -y_N) \in \bar{Y}$,

$$\frac{1}{2}(b(y_1, \ldots, y_N), y_1, \ldots, y_N) + \frac{1}{2}(-a(y_1, \ldots, y_N), -y_1, \ldots, -y_N) = \left(\frac{b(y_1, \ldots, y_N) - a(y_1, \ldots, y_N)}{2}, 0, \ldots, 0\right) \in \bar{Y},$$

and

$$\left(\frac{a(y_1, \ldots, y_N) - b(y_1, \ldots, y_N)}{2}, 0, \ldots, 0\right) \in \bar{Y}.$$

Due to the definition of the functions a and b,

$$\begin{aligned} b(0, \ldots, 0) - a(0, \ldots, 0) &\geq \frac{b(y_1, \ldots, y_N) - a(y_1, \ldots, y_N)}{2} \\ &\quad - \frac{a(y_1, \ldots, y_N) - b(y_1, \ldots, y_N)}{2} \\ &= b(y_1, \ldots, y_N) - a(y_1, \ldots, y_N). \end{aligned}$$

Now, recalling the definition of d in (1) and recalling that $(y_1, \ldots, y_N) \in \pi(Y)$ is arbitrary and Y is balanced, we get successively $b(0, \ldots, 0) - a(0, \ldots, 0) = 2d$, $b(0, \ldots, 0) = -a(0, \ldots, 0) = d$.

Assuming $y_1 = \cdots = y_N = 0$ in (2), we see that $|y_0 - p_0| \leq d$ for any

$$y_0 \in (a(0, \ldots, 0), b(0, \ldots, 0)) = (-d, d),$$

which is possible only for $p_0 = 0$. ■

In Fig. 1 the graph of the function $y_0 = p_0 + p_1 y_1 + \cdots + p_N y_N$, defining the optimal affine method approximating $S(f)$, is exhibited. Obviously any function $y_0 = \theta_0(y_1, \ldots, y_N)$, whose graph lies entirely within the shaded area in Fig. 1, defines an optimal method since $|y_0 - \theta_0(y_1, \ldots, y_N)| \leq d$ for any $(y_0, y_1, \ldots, y_N) \in Y$.

REFERENCES

BAKHVALOV, N. S. (1971), On the optimality of linear methods for operator approximation in convex classes of functions, *Zh. Vychisl. Mat. i Mat. Fiz.* **11,** 1014–1018 [Russian]; *U.S.S.R. Comput. Math. and Math. Phys.* **11,** 244–249 [English transl.].

MARCHUK, A. G., AND OSIPENKO, K. YU. (1975), Best approximation of functions specified with an error at a finite number of points, *Mat. Zametki* **17,** 359–368 [Russian]; *Math. Notes* **17,** 207–212 [English transl.].

MICCHELLI, C. A., AND RIVLIN, T. J. (1977), A survey of optimal recovery, *in* "Optimal Estimation in Approximation Theory" (C. A. Micchelli and T. J. Rivlin, Eds.), pp. 1–54, Plenum, New York.

OSIPENKO, K. YU (1976), Best approximation of analytic functions from information about their values at a finite number of points, *Mat. Zametki* **19,** 29–40 [Russian]; *Math. Notes* **19,** 17–23 [English transl.].

PACKEL, E. W. (1984), "Linear Problems (with Extended Range) Have Linear Optimal Algorithms," Report, Columbia University; *Aequationes Math.*, in press.

ROCKAFELLAR, R. T. (1970), "Convex Analysis," Princeton Univ. Press, Princeton, N. J.

SMOLYAK, S. A. (1965), "On Optimal Restoration of Functions and Functionals of Them," Candidate Dissertation, Moscow State University. [Russian]

TRAUB, J. F., AND WOŹNIAKOWSKI, H. (1980), "A General Theory of Optimal Algorithms," Academic Press, New York.

WERSCHULZ, A. G., AND WOŹNIAKOWSKI, H. (1985), "Are Linear Algorithms Always Good for Linear Problems?" Report, Columbia University; *Aequationes Math.*, in press.

WOŹNIAKOWSKI, H. (1985), A survey of information-based complexity, *J. Complexity* **1,** 11–44.